



ELSEVIER

SCIENCE @ DIRECT®

See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex

Michael S Beauchamp

Our understanding of multisensory integration has advanced because of recent functional neuroimaging studies of three areas in human lateral occipito-temporal cortex: superior temporal sulcus, area LO and area MT (V5). Superior temporal sulcus is activated strongly in response to meaningful auditory and visual stimuli, but responses to tactile stimuli have not been well studied. Area LO shows strong activation in response to both visual and tactile shape information, but not to auditory representations of objects. Area MT, an important region for processing visual motion, also shows weak activation in response to tactile motion, and a signal that drops below resting baseline in response to auditory motion. Within superior temporal sulcus, a patchy organization of regions is activated in response to auditory, visual and multisensory stimuli. This organization appears similar to that observed in polysensory areas in macaque superior temporal sulcus, suggesting that it is an anatomical substrate for multisensory integration. A patchy organization might also be a neural mechanism for integrating disparate representations within individual sensory modalities, such as representations of visual form and visual motion.

Addresses

Laboratory of Brain and Cognition, National Institute of Mental Health Intramural Research Program, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, USA

Corresponding author: Beauchamp, Michael S (mbeauchamp@nih.gov)

Current Opinion in Neurobiology 2005, 15:1–9

This review comes from a themed issue on
Cognitive neuroscience
Edited by Angela D Friederici and Leslie G Ungerleider

0959-4388/\$ – see front matter
© 2005 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.conb.2005.03.011

Introduction

In everyday life, perceptual events often occur in multiple sensory modalities at once: we hear someone speaking as we see their mouth move. Most scientific investigations have focused on single modalities (frequently vision) in isolation. Recently, there has been increasing interest in studying integration across sensory modalities. In this review, I discuss progress in studying the brain mechanisms of multisensory integration in human lateral occipital-temporal cortex, especially functional magnetic resonance imaging (fMRI) studies of superior temporal

sulcus (STS), area LO and area MT (see glossary for a brief definition of these terms). Links between human neuroimaging studies and studies in non-human primates are made using techniques from computational neuroanatomy that permit alignment of human and monkey brains.

An ongoing discussion concerns the appropriate methods for studying multisensory integration using fMRI [1*,2*,3]. One important method is to contrast unisensory stimulation conditions with multisensory conditions. The hallmark of multisensory integration is that unisensory stimuli presented in combination produce an effect different from the linear combination of the unisensory stimuli presented separately. In individual neurons, these differences can be quite dramatic, with multisensory responses that are much greater than the sum of individual unisensory responses ('super-additivity'). However, because fMRI measurements integrate across thousands or millions of neurons, the super-additivity measure might not be appropriate [2*]. Instead, increasingly liberal criteria might be more suitable, such as requiring only that multisensory responses are greater than the maximum or mean of the individual unisensory responses [1*].

Another important issue is the high degree of inter-subject and -laboratory variability observed in fMRI studies. STS, LO and MT are attractive targets for a review because there is some consensus on their anatomical location. This is either because they constitute an anatomical structure observed in every normal human hemisphere (such as STS) or because their response properties make it possible to identify them with functional localizers (somewhat ambiguously for LO, unambiguously for MT). By starting out with well-defined regions, a review can sidestep some of the difficulties inherent in deciding if a stereotaxic coordinate reported in one study of multisensory integration corresponds to the same cortical region as a coordinate from a different study.

Although STS, LO and MT are found in relative proximity, within the space of a few centimeters in human lateral occipital temporal cortex, their multisensory response properties are quite different, as is our level of knowledge about their role in multisensory perception. Therefore, this review attempts to compare and contrast the activity in these three areas in response to stimuli in three sensory modalities — visual, auditory and tactile. **Figure 1** illustrates the location of STS, LO and MT in folded and inflated versions of a human brain, and their relationship to Brodmann's cytoarchitectonic classification scheme.

Glossary

Area MT (V5): A region in extrastriate visual cortex distinguished by its heavy myelination and specialization for processing visual motion. It was first described in the posterior middle temporal cortex of owl monkey [53], leading to the designation MT. In macaque monkeys, this region lies in the posterior bank of the superior temporal sulcus, where some investigators have designated it V5 [54]. A homologous region has been found in many other species, including humans, where it lies near the junction of the inferior temporal sulcus and the lateral occipital sulcus [55].

Congruent and incongruent stimuli: Because different sensory modalities can be stimulated independently in an experimental setting, multisensory stimuli can be congruent (such as a picture of a car presented with the sound of a car) or incongruent (such as a picture of a car presented with the sound of a telephone).

fMRI (functional magnetic resonance imaging): A non-invasive method for measuring neuronal activity, typically with an indirect measure such as blood-oxygenation level dependent (BOLD) contrast.

Localizer: There is only a rough correlation between visible anatomical structures (such as specific sulci or gyri) and the functional areas that comprise the computational organization of the brain. However, in order to make inferences about organization, it is important to compare the same functional area across subjects. A common technique is to use a localizer fMRI scan (for instance, alternating moving and static stimuli) in order to identify a specific region of interest (for instance, area MT). Additional experiments are then performed and the results compared across subjects within this region.

Multisensory: Refers to the processing of stimuli presented in multiple sensory modalities at once. Although the term 'multimodal' is sometimes used as a synonym for multisensory, it is also used to describe studies that use multiple measurement techniques, such as fMRI and magnetoencephalography (MEG). Therefore, the term multisensory is preferred.

Synchronous and asynchronous stimuli: An experimental manipulation that involves artificially changing the temporal offset between stimuli presented in different sensory modalities in order to measure the effect on multisensory integration. For instance, the disconcerting sensation when the dialogue in the sound track of a movie is offset from the images.

Multisensory integration in superior temporal sulcus

There is compelling evidence for auditory and visual responses in human STS to a variety of stimuli. (For a review of all regions important for multisensory identification and object recognition, please see Amedi *et al.* [4]). Because it extends over a large area of cortex, STS certainly contains several functional regions. However, the parcellation of human STS is poorly understood, and in this review STS is used as shorthand for 'the constellation of cortical areas with multisensory response properties in STS'.

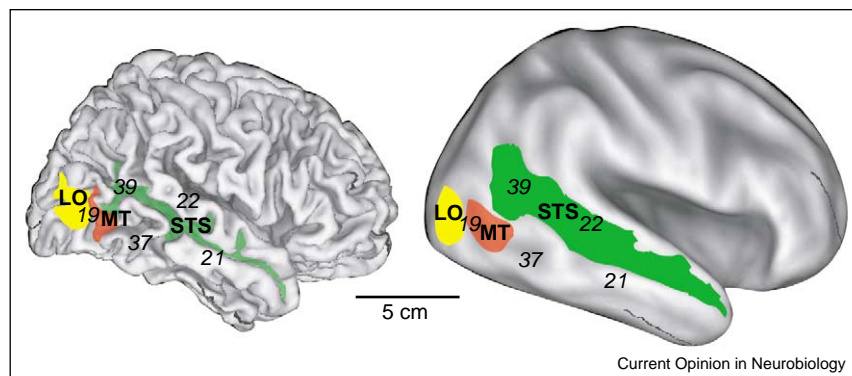
Although STS responds to simple stimuli (such as moving visual gratings) it shows a much greater response to meaningful stimuli, such as moving people or objects [5]. Speech and language related processing is one domain of meaningful stimuli in which auditory-visual integration is particularly important. Wright *et al.* [6•] measured activity in response to auditory (A), visual (V) and auditory-visual (AV) animated characters speak-

ing single words. Activity in response to the visual stimuli was strongest in the posterior half of the STS, whereas anterior regions of the STS were activated only by A and AV stimuli. AV stimuli elicited the greatest response throughout the STS, especially in mid-STS. Using simpler linguistic stimuli, Van Atteveldt *et al.* [7•] found regions in STS that responded to visually presented letters (V), spoken single letters (A), or the combination (AV). As in the study by Wright *et al.*, the STS response was greatest for AV stimuli. Interestingly, this multisensory enhancement was seen for congruent stimuli (i.e. visual 'b' and auditory 'bah') but not incongruent stimuli (e.g. 'b' and 'kah'). Because no behavioral task was required of the subjects in the Wright and Van Atteveldt studies, the enhanced response during AV stimulation (or congruent stimuli in the Van Atteveldt study) could be partially explained by increased attention and arousal. However, a similar result was found in a study of synchronous versus asynchronous AV speech in which a behavioral task was used [8], and in a study combining audiovisual objects with a behavioral task.

Multisensory responses in STS are not restricted to linguistic stimuli. Beauchamp *et al.* [9•] examined responses to stimuli representing animals and man-made graspable objects (tools). In the first experiment, subjects were presented with static pictures or sound recordings of animals and tools, in addition to scrambled auditory or visual control stimuli, and performed a one-back same-or-different task. STS showed strong activity in response to the meaningful animal or tool stimuli in both auditory and visual modalities that was greater than the level of activity in response to scrambled auditory or visual control stimuli, even though the task was equally difficult for meaningful and meaningless stimuli. In the second experiment, auditory and visual stimuli were presented in combination, and an enhanced response was observed in STS for multisensory compared with unisensory stimulation. In the third experiment, an event-related design was used in which the presentation of the sensory stimulus, consisting of videos of moving tools (V), recordings of the same tools (A), or the combination (AV), was separated in time from the behavioral decision made by the subject (selecting the name of the tool from three choices). This enabled cortical areas to be categorized into two groups. One group of areas, including STS, responded more to the sensory stimulation than to the behavioral decision, whereas another group of areas, in parietal and frontal cortex, responded more to the behavioral decision. As in the second experiment, STS showed enhanced activity during AV stimulation compared with that during A or V stimulus presentation.

To localize the anatomical location of this multisensory activity more precisely within the STS, a new analysis of the data from Beauchamp *et al.* [9•] was conducted for this review using a novel intersubject averaging technique

Figure 1



The locations of the three multisensory regions described in the review shown on a lateral view of a folded (left) and inflated (right) right hemisphere: area LO is yellow, area MT is red and STS is green. Numbers indicate approximate centers of nearby Brodmann areas. The region labeled LO corresponds retinotopically to dorsal V4 [26] and is a subset of a much larger band of cortex (extending superiorly and inferiorly) that responds preferentially to real visual images versus scrambled controls. Data can be downloaded or viewed from <http://sumsdb.wustl.edu:8081/sums/directory.do?dirid=6135667>. See Van Essen and Orban *et al.* [15[•],47] for more details.

that enables group calculations on the cortical surface [10]. As shown in Figure 2a, a bilateral region in left and right posterior STS, extending inferiorly into middle temporal gyrus, met criteria as a site for multisensory integration. The location of the active cortex within STS is consistent with other studies of auditory–visual integration.

Multisensory responses in human and monkey superior temporal sulcus

In macaque, an important multisensory region lies along the fundus of the STS. This region was functionally defined as the superior temporal polysensory (STP) area on the basis of single cell recordings [11] and probably corresponds to the region in macaques that was anatomically defined as temporal–parietal–occipital (TPO) [12]. Although the visual responses of many areas in macaque STS have been characterized, recent neuroimaging studies in macaque demonstrate that complex, behaviorally relevant sounds, such as monkey calls, also evoke activity in STS [13,14[•]]. Figure 2b shows a macaque monkey brain with the location of TPO along the floor of the STS illustrated.

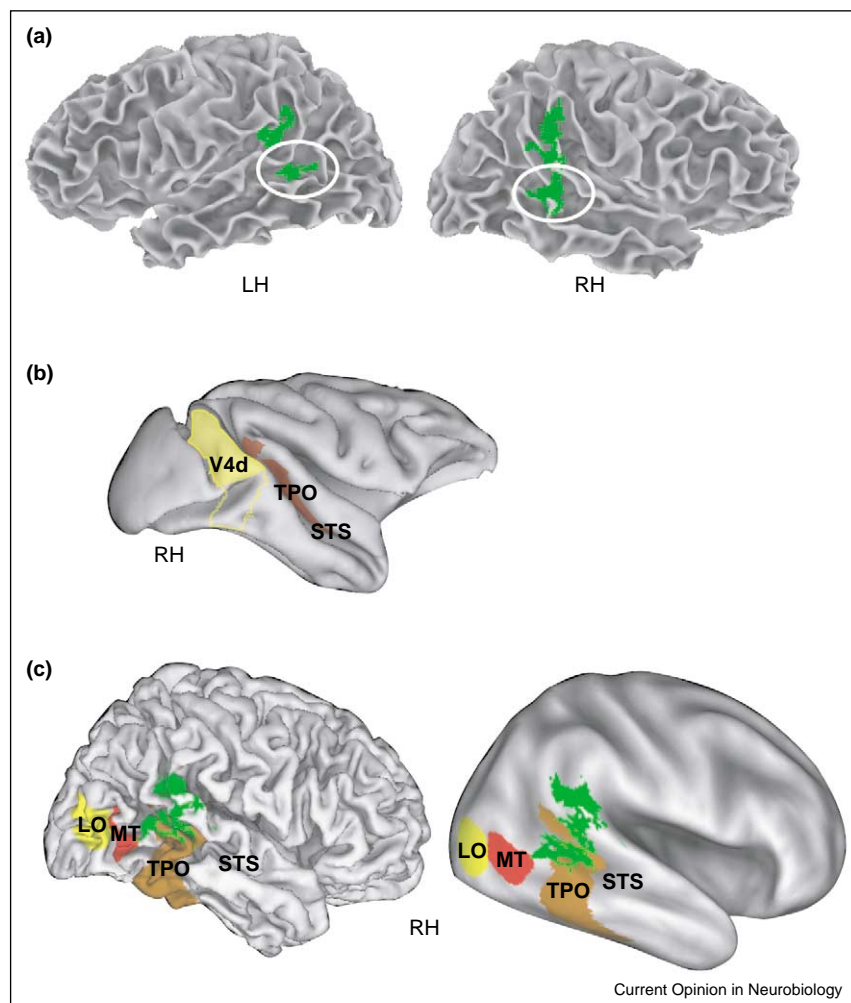
Recently developed techniques make it possible to align monkey and human brains [15[•]]. This enables predictions to be made about human cortical areas from invasive anatomical and physiological studies in non-human primates that are unavailable in humans. With an alignment scheme that uses both anatomical and functional landmarks, macaque TPO is predicted in humans to extend inferiorly from the posterior STS into middle temporal gyrus (Figure 2c; [15[•]]). But why is an area that extends anterior-to-posterior along the fundus of STS in macaques predicted to run inferiorly from STS into MTG in humans? Whereas macaque MT is located in the lower

bank of the STS, human MT is located posterior and inferior to the STS in lateral occipital temporal cortex. This suggests a differential expansion of this part of cortex in humans compared with monkeys since both species diverged from their last common ancestor. Using the relative position of functional landmarks, such as area MT, to drive the macaque–human alignment process pulls other regions, such as TPO, out of their location in the fundus of the STS of macaques and into a more posterior and inferior location in humans. This proposal for the location of a possible TPO is consistent with the human fMRI data. As shown in Figure 2c, there is substantial overlap between human multisensory activity in STS (from the reanalysis of [9[•]]) and the proposed TPO homolog.

Whereas tactile and visual integration in macaque TPO spurred the original designation of the area as polysensory, to date there is no evidence for multisensory tactile integration in human STS. STS activity has been reported in group-averaged maps from unisensory tactile studies [16], but the relative spatial position, consistency and amplitude of tactile compared with visual or auditory responses, and the optimal tactile stimuli for STS are unknown. To provide further evidence of homology between monkey TPO and human STS, it will be important to investigate the auditory, visual and tactile multisensory responses of human STS.

Studies of anatomical organization in macaques have shown that TPO comprises multiple chemoarchitectonic modules [12]. Neighboring regions of TPO receive input from auditory or visual areas in a non-overlapping fashion [17,18]. Recently, a potential functional counterpart to this patchy anatomical organization has been described in human STS (Figure 3; [19[•]]).

Figure 2

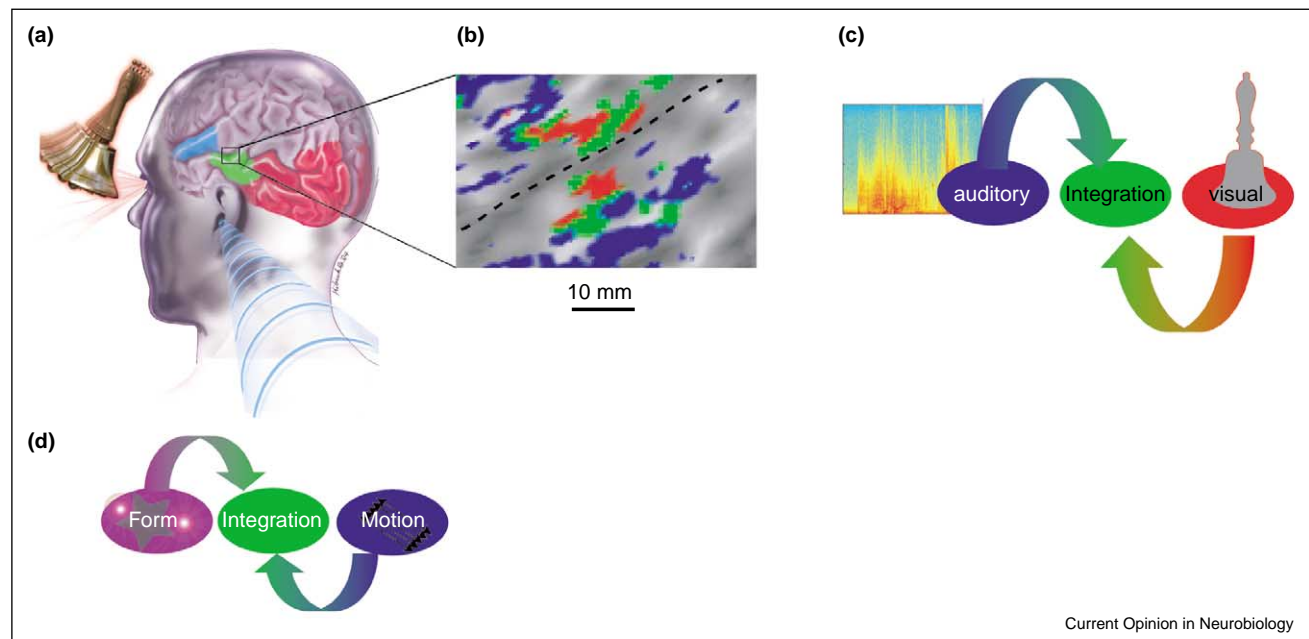


Anatomical location of multisensory activity in human and macaque monkey STS (LH indicates left hemisphere, RH indicates right hemisphere). (a) Cortical surface models were constructed from each of the eight subjects that performed experiment #3 in Beauchamp *et al.* [9]. These surface models were subjected to a spherical averaging technique in which the sulcal and gyral folding patterns of each subject were used to drive the intersubject anatomical alignment. Functional data was then mapped to the surface from each subject, and an intersubject analysis of variance (ANOVA) was performed at each surface node [48–50]. The ANOVA results were filtered to show only regions that responded to both unisensory auditory and visual stimuli (sound recordings and videos of moving tools), showed an enhanced response to multisensory (AV) stimuli, responded more during sensory stimulation than during response selection and motor output (in which subjects selected the name of the tool that they had just observed), and had an activation area on the cortical surface of $>100 \text{ mm}^2$. Multisensory area in the STS is shown with a white circle. Additional activation was also observed more superiorly, in the region of posterior Sylvian fissure at the parietal–temporal boundary, which might reflect the requirements of the naming task [51]. (b) Lateral view of a macaque brain. Yellow filled region corresponds to the average location of dorsal V4 (a possible macaque homolog of human area LO [26]), and the yellow outline indicates the location of ventral V4. Brown shaded region labeled 'TPO' indicates the approximate location of macaque polysensory cortex (areas TPOc, TPOi, and TPOr). Area MT (not visible) lies in the posterior bank of the STS. Location of areas from Lewis and Van Essen [52]. (c) Possible homologies between macaque and human STS multisensory cortex shown on folded and inflated human brains. Brown shaded region indicates the position of monkey area TPO morphed to human brain using anatomical and functional landmarks from Denys *et al.* [24]. Green shaded region illustrates location of STS multisensory activity measured with fMRI, from (a). Data can be viewed or downloaded from <http://sumsdb.wustl.edu:8081/sums/archivist.do?dirid=6135667>.

Multisensory cortex in the STS was localized using standard-resolution fMRI while subjects were presented with sound recordings and videos of moving tools. Then, in a separate scanning session, STS multisensory cortex was examined at much higher resolution using parallel imaging. By combining signals from multiple receiver coils,

each of which is most sensitive to a small region of tissue, parallel imaging fMRI [20,21] enables measurement of the blood-oxygenation signal at an order of magnitude higher in resolution (voxel size $\sim 1\text{--}5 \text{ mm}^3$) than standard resolution fMRI (voxel size $\sim 50\text{--}100 \text{ mm}^3$). Because a much smaller total brain volume can be scanned at this

Figure 3



A schematic model of how patchy organization in STS might subserve multisensory integration. (a) Sensory input in the auditory and visual sensory modalities from objects (such as a bell, shown) is processed by auditory (blue) and visual (orange) cortex. STS (green) integrates auditory and visual information. (b) STS multisensory cortex (fundus of STS shown with black dashed line) contains patches of cortex that respond preferentially to auditory stimuli (blue), visual stimuli (orange) or both auditory and visual stimuli (green). (c) Within STS, anatomical data suggest that auditory patches (single patch shown schematically in blue) receive input from auditory association cortex, which uses a code focused on acoustic spectral frequency content (spectral frequency plot of bell sound shown). Visual patches (orange) receive input from visual association cortex, which encodes higher level object features, such as shape (shape of bell shown). Local connections between patches might serve to convert visual and auditory codes to a common code that can then be used for multisensory integration (green patch). (d) More speculatively, similar mechanisms might underlie the integration of visual form and motion in STS. Patches might receive input from form and motion selective regions of visual cortex, followed by integration in the intervening patches.

higher resolution, the initial localizer scans were necessary to target the precise location of STS multisensory cortex in each individual subject.

At standard resolution, STS multisensory cortex appeared uniformly sensitive to both auditory and visual stimuli (Figure 3a). At high resolution, the activity resolved into discrete patches that responded primarily to auditory, primarily to visual, or to both auditory and visual stimuli (Figure 3b). Only the patches that responded to both auditory and visual stimuli showed the enhanced response to multisensory stimuli that is the hallmark of multisensory integration. The same organization was observed for two completely different stimulus sets: videos of tools (V), recordings of tools (A), or the combination (AV); and videos of faces (V), recordings of voices (A), or the combination (AV). This supports the idea that the patchy structure observed with fMRI reflects patchy visual and auditory inputs into multisensory cortex, as found in anatomical studies of macaques (discussed above). The patchy organization might be the result of a processing strategy in which visual and auditory inputs arrive in neighboring, but separate, regions of STS multisensory

cortex, and are then integrated in the intervening patches that respond to both auditory and visual stimuli (Figure 3c).

From an information processing standpoint, what would be the reason for this organization? Auditory and visual inputs to STS are coded in very different dimensions. Auditory association areas that project to STS code in terms of spectral content, whereas visual areas that project to STS code in terms of form and motion. One possibility is that auditory and visual patches in STS take these very different representations and convert them into a common code, which is then passed via local connections to the intervening multisensory patches, where the auditory–visual representations are integrated and sent forward to higher-level processing centers (Figure 2c).

Area LO

Area LO was first described as a region of human lateral occipital cortex, just ventral and posterior to area MT, that responded preferentially to images of objects versus those of textured patterns [22]. LO is thought to be important for processing visual shape information [23]. More

recently, studies showing that an extended band of visual cortex responds preferentially to images versus patterns [24*,25] has led to confusion over the location and identity of LO. Figure 2 illustrates the location of human dorsal V4, which comprises part of this large area of image-preferring cortex and might correspond to LO [26]. Although little is known about multisensory responses in monkey dorsal V4, human fMRI studies have suggested that tactile shape information, as well as visual shape information, is processed in area LO. Amedi *et al.* [27] required subjects to identify real objects (e.g. forks) and textures (e.g. sandpaper) by touch. LO was localized with a contrast between visual and scrambled objects. Portions of visually defined LO, named 'LOtv' (tactile-visual) by the authors, showed stronger activity in response to objects identified by touch than to touched textures (such as sandpaper), in contrast to somatosensory cortex, which responded similarly to the two conditions. James *et al.* [28] found similar results using novel objects made out of clay. In the James study, although an LO localizer was not performed, both visual and tactile exploration of novel objects activated regions posterior to MT, both superiorly (in a region which the authors refer to as 'MO', for its location on the middle occipital gyrus) and inferiorly (which the authors refer to as 'LO'). Cross-modality priming effects were observed, suggesting that LO was engaged in representing a higher-order shape representation accessible by vision or touch. Pietrini *et al.* [29] also studied activity in response to tactile and visual identification of real objects. A region with the same standardized coordinates as LO (which the authors refer to as 'inferior temporal', or IT) responded to identification of both tactile and visual objects. Other studies that have examined purely tactile form discrimination or object recognition (without a visual counterpart) have also observed activation in LO-like regions [30–33].

An obvious question is the role of visual imagery in the observed tactile LO responses. It is important to distinguish two axes of concern. One axis concerns the origins of LO activity, either from a bottom-up source (ascending inputs from somatosensory cortex) or from a more indirect route via visual imagery. Amedi *et al.* [27] addressed this concern by asking subjects to create mental images of the objects that they had identified visually and haptically. Although some activation in LO in response to imagery was observed, this activation was about four times less than that evoked by actual tactile object recognition. By contrast, Zhang *et al.* [30] found a high correlation (0.90) between activity in right LO during tactile form perception and an independent measure of subjects' visual imagery obtained via questionnaire and self-report. Although the importance of visual imagery in generating tactile activity in LO in normal subjects is unclear, a more fundamental axis of concern is whether the activity in LO has functional relevance for tactile form recognition or is simply epiphenomenal. Convincing evidence of its

importance comes from studies of patient DF, who has large bilateral lesions that include LO [34], and is impaired at learning novel objects presented haptically [28]. Similarly, Pietrini *et al.* [29] observed LO responses during haptic object recognition in congenitally blind subjects, for whom visual imagery is unavailable. These results show that tactile stimuli can evoke LO activity directly (bypassing imagery) and seem to confirm that LO is a necessary part of the neural substrate for processing shape information, whether derived from visual or tactile sources.

Area MT

Area MT is recognized as a key locus for visual motion processing in the primate brain (see glossary). In macaque monkeys, MT is located in the lower bank of the STS (Figure 2b), whereas in humans, MT is located in lateral occipital cortex (Figure 2c). This review refers to 'MT' as a single area for simplicity, although this region of cortex contains several motion-responsive areas that are grouped together in most imaging studies, often under the rubric MT+ [35].

There are strong interconnections between monkey ventral intraparietal area (VIP), a parietal region sensitive to tactile stimuli, and MT, providing a possible anatomical route for tactile information to reach MT. Although tactile responses have not been reported in monkey MT, two recent human functional neuroimaging studies suggest that MT might have a role in processing motion in the tactile modality. Hagen *et al.* [36] found greater activity in area MT when a small brush stroked the length of the subject's arms than that during fixation control. Blake *et al.* [37] found greater activity in area MT as subjects grasped a rotating plastic ball than that when they grasped a stationary ball. As discussed above for LO, one might ask if tactile activity in MT arises from bottom-up sensory input or top-down cognitive strategy, such as imagery. Psychophysically, when subjects in the study by Blake *et al.* [37] were adapted to the rotating tactile globe, it did not affect their visual judgment of globe direction, in contrast to the adaptation effects observed with visual adaptation (and what would be expected if visual imagery were the cause of MT activity). Subjects in the Blake *et al.* [37] study were explicitly asked to imagine rotation of the globe, and activity in area MT was not observed. However, this finding conflicts with that from Goebel *et al.* [38], which described MT activity during motion imagery.

Regardless of the source of tactile responses in MT, they are relatively weak. In the LO studies, activity in parts of LO was nearly as great for tactile as for visual stimuli, whereas in STS auditory activations are as large or larger than visual activations. By contrast, the reported MT activations are much weaker for moving tactile than moving visual stimuli (0.8% MR signal change versus 0.2% in the Blake study). Although these tactile responses are of low amplitude, they are positive, in contrast to the

negative signal changes (below fixation baseline) observed in MT when subjects performed an auditory motion discrimination task [39], a non-motion tactile task [16] and a non-motion auditory speech task [6^{*}]. Therefore, the weak tactile motion response in area MT might be the result of a conflict between activation and deactivation [40]. In addition, weak responses in MT to tactile motion stimuli might simply mirror the fact that for motion processing in primates, the visual modality is dominant over other modalities. This effect was quantified by Soto-Faraco *et al.* [41], who showed that a visual distractor moving in one direction could ‘capture’ auditory or tactile target stimuli moving in the opposite direction, and cause them to be perceived as moving in the same direction as the visual moving stimulus (even though subjects were instructed to ignore the visual stimulus). Tactile moving stimuli could also capture auditory stimuli, whereas only rarely could auditory stimuli capture tactile stimuli, and auditory and tactile stimuli never captured visual stimuli. It is tempting to speculate that the results might reflect the connectedness of MT to areas that process tactile stimuli, such as VIP, in relation to the lack of connectedness of MT to auditory inputs, such as belt or parabelt regions of auditory cortex.

As mentioned in the Introduction, a crucial test of multisensory integration is the difference between unisensory and multisensory responses. Unlike in STS and LO, these measurements have not yet been made in area MT. Such experiments will certainly shed more light on the relative influence of task factors, such as imagery, on tactile activity in area MT.

Commonalities between integration across and within modalities

As discussed above, one of the neural substrates for multisensory integration in STS might be a patchy organization, in which neighboring patches respond primarily to unisensory auditory or visual information. Unisensory information might be translated into a common code and integrated in multisensory regions that lie between the unisensory patches. Such an organization might also be amenable to integration of other types of information. Neurons in STS can be selective to both visual form and visual motion, for instance responding only when the back view of an individual is observed moving away from the subject [42]. Because form and motion are processed in different visual areas, these visual primitives might also arrive in STS in different patches, followed by integration in intervening cortex (Figure 3d). It will be important to investigate the possibility of patchy inputs into STS for inputs from within visual submodalities (such as form and motion). Evidence supports a patchy visual–tactile organization in LO. When examined in single subjects, tactile responses in LO do not extend over the entire region, but instead are concentrated in several relatively small, discrete regions [43].

The object property model

It is also useful to consider the relationship between multisensory and category-related responses. One of the most surprising findings to arise from recent functional neuroimaging studies is that specific regions of human visual cortex respond preferentially to specific categories of objects. For instance, parts of lateral temporal cortex (middle temporal gyrus and inferior temporal sulcus, including portions of areas MT and LO) respond preferentially to images of man-made graspable objects (tools) versus images of other categories of objects. These regions of lateral temporal cortex also respond to visual motion [5]. The object property model hypothesizes that because a crucial identifying property of individual tools is their motion (e.g. the characteristic up-and-down movement of a hammer) the neural representation of tools is linked to visual motion-responsive cortex in lateral temporal cortex [44]. A similar argument can be made for tactile responses to objects [43]. Because tools have familiar tactile properties (unlike other categories of objects with distinct neural representations, such as houses) the object property model predicts that tactile responses in lateral temporal cortex should be concentrated within regions that respond preferentially to visually presented tools. This was found in the study by Amedi *et al.* [43], who found that tactile responses in LO overlapped regions that responded to visually presented tools but not regions that responded preferentially to visually presented faces or houses. Consistent with this finding, Pietrini *et al.* [29] found that cortex that responded to visually presented bottles or shoes also responded to haptic inspection of bottles and shoes, and showed similar patterns of activities across modalities [29]. However, this effect was not seen for visual and tactile inspection of faces, which are not normally recognized in the tactile domain. These findings provide support for the object property model by showing that tactile activity in LO is evoked only by objects for which tactile information is important (bottles, shoes and other graspable objects).

It will be important to test the object property model in the STS, where it makes clear predictions. Studies of visual presentation of biological motion stimuli show that different regions of STS are preferentially active for hand, mouth and eye movements [45,46]. Regions in STS specialized for processing visual mouth movements should be more sensitive to auditory cues than region that respond preferentially to eye-movement, because of the greater association between mouth movements and auditory stimulation.

Limitations of the object property model are apparent when we consider responses to auditory stimuli in MT and LO. Whereas auditory information does not definitively specify object shape, hearing the ‘ring-ring’ of a telephone or the ‘bang-bang’ of a hammer gives some

information about the shape that the object is likely to be. However, auditory object representations produce no activation in area LO [43]. Moving auditory stimuli provide information about probable visual motion, but auditory motion in isolation actually reduces activity in area MT below baseline [39]. These findings are not surprising if we consider the connectivity of these regions. Whereas area MT is strongly interconnected with area VIP, which receives both visual and tactile input, there is little evidence of similar auditory pathways into area MT. Although the object property model has power to predict and explain the representation of different object categories in different brain regions, the anatomical and functional connections among regions are also an important determinant of cortical representations.

Conclusions and future directions

Neuroimaging studies in humans and non-human primates using the same multisensory stimuli will be crucial for forming a link between human neurobiology and the anatomical and physiological insight that can only be obtained from invasive studies. The results of these experiments, combined with advances in neuroimaging methods applicable in humans, such as high-resolution fMRI and MEG, mean that the next few years will surely see further great strides in our understanding of multisensory integration.

Acknowledgements

D Van Essen, D Hanlon, J Dickson, P Christidis and Z Saad were instrumental in figure preparation. A Martin and A Amedi provided helpful comments on the manuscript. This work was supported by the National Institute of Mental Health Intramural Research Program.

References and recommended reading

Papers of particular interest, republished within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Beauchamp MS: **Statistical criteria in fMRI studies of multisensory integration.** *Neuroinformatics* 2005. In press. Although fMRI studies of multisensory integration use complex statistical criteria to classify whether areas of activation are multisensory or not, there is no general agreement about the correct criteria to use. This paper is structured as an illustrated tutorial by examining the results of applying commonly used criteria to the same test dataset. Both single subject analysis (using cortical surface models to illustrate whole-brain activation patterns) and intersubject averaging methods are discussed.
2. Laurienti PJ, Perrault TJ Jr, Stanford TR, Wallace MT, Stein BE: **On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies.** *Experimental Brain Research* 2005. In press. In an effort to better understand the observed blood oxygenated level-dependent (BOLD) signal in fMRI studies of multisensory integration, the authors construct a model formed on the basis of single neuron responses properties in the superior colliculus. Although some single neurons display superadditive responses (i.e. audiovisual > auditory + visual) the BOLD signal is found to be subadditive (i.e. audiovisual < auditory + visual).
3. Calvert GA, Thesen T: **Multisensory integration: methodological approaches and emerging principles in the human brain.** *J Physiol (Paris)* 2004, **98**:191-205.
4. Amedi A, Kriegstein KV, Van Atteveldt N, Beauchamp MS, Naumer M: **Functional imaging of human crossmodal identification and object recognition.** *Experimental Brain Research* 2005. In press.
5. Beauchamp MS, Lee KE, Haxby JV, Martin A: **Parallel visual motion processing streams for manipulable objects and human movements.** *Neuron* 2002, **34**:149-159.
6. Wright TM, Pelphrey KA, Allison T, McKeown MJ, McCarthy G: **Polysensory interactions along lateral temporal regions evoked by audiovisual speech.** *Cereb Cortex* 2003, **13**:1034-1043. The authors examine brain activity in response to animated figures speaking whole words. They report auditory and visual responses along the entire length of the STS, with varying degrees of multisensory integration, suggesting possible functional subdivisions within human STS.
7. Van Atteveldt N, Formisano E, Goebel R, Blomert L: **Integration of letters and speech sounds in the human brain.** *Neuron* 2004, **43**:271-282. The authors use fMRI to measure responses to letters and their corresponding speech sounds. Interestingly, some anterior STS regions respond to bimodal stimuli but do not respond to unimodal stimuli.
8. Macaluso E, George N, Dolan R, Spence C, Driver J: **Spatial and temporal factors during processing of audiovisual speech: a PET study.** *Neuroimage* 2004, **21**:725-732.
9. Beauchamp MS, Lee KE, Argall BD, Martin A: **Integration of auditory and visual information about objects in superior temporal sulcus.** *Neuron* 2004, **41**:809-823. In three separate experiments, posterior STS and MTG were shown to respond to both auditory and visual representations of objects with greater activity than that in response to scrambled controls, and to show the greatest activation response to combined auditory-visual objects. Considerable intersubject variability is observed in the precise location of STS multisensory activity. Different behavioral tasks modulate the strength of multisensory activity, especially in frontal regions.
10. Argall BD, Saad ZS, Beauchamp MS: **A simplified method for intersubject averaging on the cortical surface using SUMA.** *Human Brain Mapping* in press.
11. Bruce C, Desimone R, Gross CG: **Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque.** *J Neurophysiol* 1981, **46**:369-384.
12. Padberg J, Seltzer B, Cusick CG: **Architectonics and cortical connections of the upper bank of the superior temporal sulcus in the rhesus monkey: an analysis in the tangential plane.** *J Comp Neurol* 2003, **467**:418-434.
13. Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L, Mishkin M: **Functional mapping of the primate auditory system.** *Science* 2003, **299**:568-572.
14. Gil-da-Costa R, Braun A, Lopes M, Hauser MD, Carson RE, Herscovitch P, Martin A: **Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque.** *Proc Natl Acad Sci USA* 2004, **101**:17516-17521. An important study demonstrating the possibility of using functional neuroimaging to examine responses to auditory stimuli in awake, behaving macaque monkeys. Socially relevant stimuli (species-specific calls) evoke activity in many temporal areas.
15. Van Essen DC: **Surface-based approaches to spatial localization and registration in primate cerebral cortex.** *Neuroimage* 2004, **23**:S97-S107. The author describes the state-of-the-art methods to compare monkey and human brains objectively.
16. Burton H, Sinclair RJ, McLaren DG: **Cortical activity to vibrotactile stimulation: an fMRI study in blind and sighted individuals.** *Hum Brain Mapp* 2004, **23**:210-218.
17. Seltzer B, Cola MG, Gutierrez C, Masee M, Weldon C, Cusick CG: **Overlapping and nonoverlapping cortical projections to cortex of the superior temporal sulcus in the rhesus monkey: double anterograde tracer studies.** *J Comp Neurol* 1996, **370**:173-190.
18. Cusick CG, Seltzer B, Cola M, Griggs E: **Chemoarchitectonics and corticocortical terminations within the superior temporal sulcus of the rhesus monkey: evidence for subdivisions of superior temporal polysensory cortex.** *J Comp Neurol* 1995, **360**:513-535.

19. Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A:
 • **Unraveling multisensory integration: patchy organization within human STS multisensory cortex.** *Nat Neurosci* 2004, **7**:1190-1192.
- High-resolution fMRI is used to explore the fine details of organization of STS multisensory cortex. Instead of the homogenous organization observed with standard resolution fMRI, auditory, visual and auditory-visual responses are concentrated in separate millimeter-scale patches of cortex.
20. De Zwart JA, Ledden PJ, Van Gelderen P, Bodurka J, Chu R, Duyn JH: **Signal to- noise ratio and parallel imaging performance of a 16-channel receive-only brain coil array at 3.0 Tesla.** *Magn Reson Med* 2004, **51**:22-26.
21. Bodurka J, Ledden PJ, Van Gelderen P, Chu R, De Zwart JA, Morris D, Duyn JH: **Scalable multichannel MRI data acquisition system.** *Magn Reson Med* 2004, **51**:165-171.
22. Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB: **Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex.** *Proc Natl Acad Sci USA* 1995, **92**:8135-8139.
23. Kourtzi Z, Kanwisher N: **Representation of perceived object shape by the human lateral occipital complex.** *Science* 2001, **293**:1506-1509.
24. Denys K, Vanduffel W, Fize D, Nelissen K, Peuskens H, Van Essen D, Orban GA: **The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study.** *J Neurosci* 2004, **24**:2551-2565.
- Comparisons of monkey and human brains are hindered by the different techniques available in the two species. The authors remedy this problem by measuring neural activity using the same technique (fMRI) the same stimuli (real and scrambled pictures) and even the same behavioral task (dimming or orientation detection).
25. Grill-Spector K, Malach R: **The human visual cortex.** *Annu Rev Neurosci* 2004, **27**:649-677.
26. Tootell RB, Hadjikhani N: **Where is 'dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence.** *Cereb Cortex* 2001, **11**:298-311.
27. Amedi A, Malach R, Hendler T, Peled S, Zohary E: **Visuo-haptic object-related activation in the ventral visual pathway.** *Nat Neurosci* 2001, **4**:324-330.
28. James TW, James KH, Humphrey GK, Goodale MA: **Do visual and tactile object representations share the same neural substrate?** In *Touch and Blindness: Psychology and Neuroscience*. Edited by: MA Heller, S Ballesteros. Lawrence Erlbaum; 2005: in press.
29. Pietrini P, Furey ML, Ricciardi E, Gobbi MI, Wu WH, Cohen L, Guazzelli M, Haxby JV: **Beyond sensory images: object-based representation in the human ventral pathway.** *Proc Natl Acad Sci USA* 2004, **101**:5658-5663.
30. Zhang M, Weisser VD, Stilla R, Prather SC, Sathian K: **Multisensory cortical processing of object shape and its relation to mental imagery.** *Cogn Affect Behav Neurosci* 2004, **4**:251-259.
31. Prather SC, Votaw JR, Sathian K: **Task-specific recruitment of dorsal and ventral visual areas during tactile perception.** *Neuropsychologia* 2004, **42**:1079-1087.
32. Stoesz MR, Zhang M, Weisser VD, Prather SC, Mao H, Sathian K: **Neural networks active during tactile form perception: common and differential activity during macrospatial and microspatial tasks.** *Int J Psychophysiol* 2003, **50**:41-49.
33. Reed CL, Shoham S, Halgren E: **Neural substrates of tactile object recognition: an fMRI study.** *Hum Brain Mapp* 2004, **21**:236-246.
34. James TW, Culham J, Humphrey GK, Milner AD, Goodale MA: **Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study.** *Brain* 2003, **126**:2463-2475.
35. Beauchamp MS, Cox RW, DeYoe EA: **Graded effects of spatial and featural attention on human area MT and associated motion processing areas.** *J Neurophysiol* 1997, **78**:516-520.
36. Hagen MC, Franzen O, McGlone F, Essick G, Dancer C, Pardo JV: **Tactile motion activates the human middle temporal/V5 (MT/V5) complex.** *Eur J Neurosci* 2002, **16**:957-964.
37. Blake R, Sobel KV, James TW: **Neural synergy between kinesthetic vision and touch.** *Psychol Sci* 2004, **15**:397-402.
38. Goebel R, Khorram-Sefat D, Muckli L, Hacker H, Singer W: **The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery.** *Eur J Neurosci* 1998, **10**:1563-1573.
39. Lewis JW, Beauchamp MS, DeYoe EA: **A comparison of visual and auditory motion processing in human cerebral cortex.** *Cereb Cortex* 2000, **10**:873-888.
40. Laurienti PJ, Burdette JH, Wallace MT, Yen YF, Field AS, Stein BE: **Deactivation of sensory-specific cortex by cross-modal stimuli.** *J Cogn Neurosci* 2002, **14**:420-429.
41. Soto-Faraco S, Kingstone A, Spence C: **Multisensory contributions to the perception of motion.** *Neuropsychologia* 2003, **41**:1847-1862.
42. Jellema T, Maassen G, Perrett DI: **Single cell integration of animate form, motion and location in the superior temporal cortex of the macaque monkey.** *Cereb Cortex* 2004, **14**:781-790.
43. Amedi A, Jacobson G, Hendler T, Malach R, Zohary E: **Convergence of visual and tactile shape processing in the human lateral occipital complex.** *Cereb Cortex* 2002, **12**:1202-1212.
44. Martin A, Chao LL: **Semantic memory and the brain: structure and processes.** *Curr Opin Neurobiol* 2001, **11**:194-201.
45. Wheaton KJ, Thompson JC, Syngieniotis A, Abbott DF, Puce A: **Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex.** *Neuroimage* 2004, **22**:277-288.
46. Hoffman EA, Haxby JV: **Distinct representations of eye gaze and identity in the distributed human neural system for face perception.** *Nat Neurosci* 2000, **3**:80-84.
47. Orban GA, Van Essen D, Vanduffel W: **Comparative mapping of higher visual areas in monkeys and humans.** *Trends Cogn Sci* 2004, **8**:315-324.
48. Argall BD, Saad ZS, Beauchamp MS: **A simplified method for intersubject averaging on the cortical surface using SUMA.** *Human Brain Mapping* 2005, in press.
49. Fischl B, Sereno MI, Tootell RB, Dale AM: **High-resolution intersubject averaging and a coordinate system for the cortical surface.** *Hum Brain Mapp* 1999, **8**:272-284.
50. Saad ZS, Reynolds RC, Argall BD, Japee S, Cox RW: **Suma: an interface for surface-based intra- and inter-subject analysis with Afni.** In *Proceedings of 2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro; Arlington, VA.* 2004: 1510-1513.
51. Hickok G, Buchsbaum B, Humphries C, Muftuler T: **Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt.** *J Cogn Neurosci* 2003, **15**:673-682.
52. Lewis JW, Van Essen DC: **Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex.** *J Comp Neurol* 2000, **428**:79-111.
53. Allman JM, Kaas JH: **Representation of the visual field in striate and adjoining cortex of the owl monkey (*Aotus trivirgatus*).** *Brain Res* 1971, **35**:89-106.
54. Zeki S: **The distribution of wavelength and orientation selective cells in different areas of monkey visual cortex.** *Proc R Soc Lond B Biol Sci* 1983, **217**:449-470.
55. Watson JD, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, Zeki S: **Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging.** *Cereb Cortex* 1993, **3**:79-94.